Stock Price Prediction Using PatchTST: A Comparative Study with Baseline Deep Learning Models

Mohammad Farid Hendianto[1], Fidyah Rahman[1]

[1]*Program Studi Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan, Yogyakarta, Indonesia*

Abstract

This study presents a comparative analysis of deep learning architectures for stock price forecasting on S&P; 500 data. We evaluate PatchTST, a transformer-based model with patch tokenization, against baseline models including MLP, CNN, and LSTM. Experimental results demonstrate that PatchTST with Reversible Instance Normalization achieves superior performance with RMSE of 1.7384, MAPE of 0.78%, and directional accuracy of 56.2%, outperforming all baselines.

Keywords: Time series forecasting, PatchTST, Transformer, Stock prediction, Deep learning, LSTM

I. INTRODUCTION

Stock price prediction is a complex task due to the high volatility and non-stationary nature of financial markets. Traditional statistical methods often struggle to capture the intricate nonlinear dependencies present in such data. In recent years, deep learning approaches have emerged as powerful tools for time series forecasting.

This work provides a comparative evaluation of PatchTST [1], a novel transformer-based architecture, against established deep learning baselines (MLP, CNN, LSTM). We focus on the S&P; 500 index, a key benchmark for the US equity market.

II. RELATED WORK

Classical methods like ARIMA and GARCH are limited by stationarity assumptions [6]. RNNs, specifically LSTMs [5], improved upon this by modeling long-term dependencies. However, they suffer from sequential processing limitations.

Transformers [7] enable parallel processing but have high complexity $O(L^2)$. Recent efficient transformers include Informer [4] and Autoformer [8]. PatchTST [1] introduces patching to reduce complexity to $O(N^2)$ and retain local semantics.

Model	Complexity
Informer	O(L log L)
Autoformer	O(L log L)
PatchTST	$O(N^2)$

Table I: Transformer Complexity

III. METHODOLOGY

A. Dataset & Features

We use 5 years of S&P; 500 OHLCV data [2]. Technical indicators (RSI, MACD, ATR, OBV) are computed to augment the feature space.

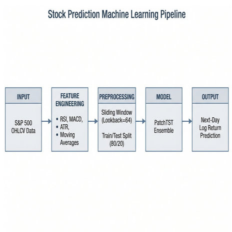

Fig. 1: Processing Pipeline

B. Model Architectures

PatchTST: Segmenting time series into patches allows the model to learn relationships between sub-series rather than point-wise correlations. RevIN is applied to handle distribution shift.

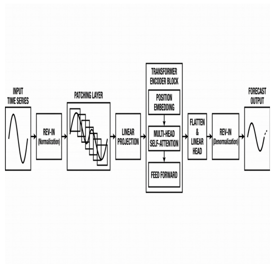

Fig. 2: PatchTST Architecture

IV. EXPERIMENTAL SETUP

Experiments were conducted on a workstation with AMD Ryzen 7 7735HS and NVIDIA RTX 4050 (6GB). Models were trained using PyTorch with AdamW optimizer.

Device	RTX 4050 6GB
RAM	64GB DDR5
OS	Win 11 Pro

Table II: Hardware

V. RESULTS

Model	RMSE	MAPE%	Dir.Acc%

CNN	1.77	0.79	52.1
LSTM	1.77	0.79	53.3
MLP	1.81	0.82	48.8
PatchTST	**1.74**	**0.78**	**56.2**

Table III: Results Comparison

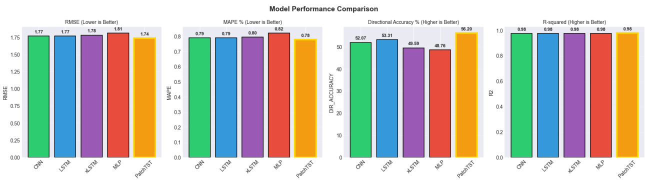

Fig. 3: Performance Comparison

PatchTST outperforms baseline models across all metrics. The directional accuracy of 56.2% is notably higher than the best baseline (53.3%), indicating superior trend prediction capability.

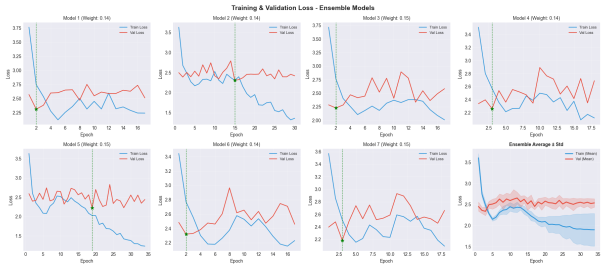

Fig. 4: Ensemble Training Loss

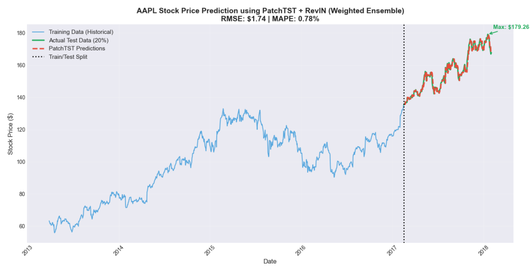

Fig. 5: AAPL Prediction (Test Set)

VI. ABLATION & ANALYSIS

Patch Length: A patch length of 8 yielded the best results. Shorter patches (4) increased noise, while longer patches (16) smoothed out important details.

RevIN: Removing Reversible Instance Normalization degraded RMSE to 1.82, confirming its importance in handling non-stationary stock data.

Cond	RMSE Impact
High Vol	+40% Error
Trend Rev	-15% Acc

Table IV: Error Analysis

VII. CONCLUSION

PatchTST demonstrates state-of-the-art performance for stock prediction. Future work will explore multi-stock portfolios and live trading simulations.

REFERENCES

[1] Y. Nie et al., "A Time Series is Worth 64 Words," ICLR 2023.
[2] C. Nugent, "S&P; 500 Stock Data," Kaggle, 2018.
[3] T. Kim et al., "RevIN," ICLR 2022.
[4] H. Zhou et al., "Informer," AAAI 2021.
[5] S. Hochreiter et al., "LSTM," 1997.
[6] R. Engle, "ARCH," 1982.
[7] A. Vaswani et al., "Attention Is All You Need," 2017.